

Implementasi Data Mining Untuk Mengelompokkan Korban Penyalahgunaan Narkoba Menggunakan Metode *K-Means Clustering* Pada BNN Provinsi Sumatera Utara

Yanuar Ishaq Hasibuan*, Puji Sari Ramadhan**, Milfa Yetri, ***
1,2,3 Program Studi Sistem Informasi, STMIK Triguna Dharma

Article Info

Article history:

Received Aug 12th, 2021

Revised Aug 20th, 2021

Accepted Aug 30th, 2021

Keyword:

Korban Penyalahgunaan Narkoba,
Data Mining,
K-Means

Pemerintah sudah menilai kejadian narkoba termasuk dalam kategori darurat narkoba. Hal ini didasarkan pada banyak kasus yang terjadi, peredaran narkoba tidak hanya pada orang-orang dewasa saja, tetapi sudah di kalangan generasi muda anak-anak Indonesia. BNN Provinsi Sumatera Utara merupakan salah satu instansi pemerintah yang menanggulangi permasalahan narkoba dan obat-obatan terlarang. Namun pihak instansi memiliki kendala terkait cara dalam mengelompokkan data penyalahgunaan narkoba tersebut, salah satu kendalanya adalah keterbatasan pengetahuan tentang metode yang akan digunakan. Untuk menyelesaikan masalah tersebut dapat menggunakan aplikasi data mining, yaitu dengan memanfaatkan data yang ada untuk menggali informasi baru. Salah satu teknik yang ada pada data mining adalah clustering. Clustering dipilih karena dapat mengelompokkan data-data sesuai dengan karakteristik yang diinginkan, dalam penelitian ini berarti mengelompokkan data korban penyalahgunaan narkoba di BNN Provinsi Sumatera Utara. Adapun algoritma clustering yang digunakan adalah *K-Means Clustering* diintegrasikan pada aplikasi pemrograman berbasis dekstop. Kesimpulan yang diperoleh yaitu sistem yang dirancang telah terbukti berhasil mengelompokkan data korban penyalahgunaan narkoba menjadi 3 cluster (kelompok). Cluster 1 merupakan cluster terkecil dengan 5 data korban, cluster 2 terdiri dari 29 data korban dan cluster 3 yang merupakan cluster terbanyak terdiri dari 26 data.

Copyright © 2021 STMIK Triguna Dharma.
All rights reserved.

ABSTRAK

Corresponding Author

Nama : Yanuar Ishaq Hasibuan

Program Studi : Sistem Informasi STMIK Triguna Dharma

Email : yanuarishaq193@gmail.com

1. PENDAHULUAN

Perkembangan kasus penyalahgunaan narkoba di Indonesia beberapa kurun waktu terakhir ini menunjukkan tingkat kenaikan yang sangat mengkhawatirkan. Jumlah pengguna narkoba di Indonesia pada tahun 2014 terhitung sebanyak 3,8 juta jiwa. Bahkan pada tahun 2015, diperkirakan jumlah pengguna narkoba di Indonesia mencapai 5,8 juta jiwa [1]. Pemerintah sudah menilai kejadian narkoba termasuk dalam kategori darurat narkoba. Hal ini didasarkan pada banyak kasus yang terjadi, peredaran narkoba tidak hanya pada orang-orang dewasa saja, tetapi sudah di kalangan generasi muda anak-anak Indonesia. BNN Provinsi Sumatera Utara merupakan salah satu instansi pemerintah yang menanggulangi permasalahan narkoba dan obat-obatan terlarang. Namun pihak instansi memiliki kendala terkait

cara dalam mengelompokkan data penyalahgunaan narkoba tersebut, salah satu kendalanya adalah keterbatasan pengetahuan tentang metode yang akan digunakan.

Untuk menyelesaikan masalah tersebut dibutuhkan suatu sistem terkomputerisasi yang dapat menggali informasi baru dari tumpukan data-data lama, yaitu dengan menggunakan *Data Mining*. *Data Mining* adalah proses yang menggunakan statistik, matematika, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [2].

Dalam penelitian ini menggunakan salah satu metode data mining yaitu *K-Means Clustering*. Algoritma *K-Means* adalah algoritma *Clustering* yang paling sederhana dibanding algoritma *Clustering* yang lain. Algoritma ini mempunyai kelebihan mudah diterapkan dan dijalankan, relatif cepat, mudah untuk diadaptasi, dan paling banyak dipraktekkan dalam tugas *Data Mining* [3].

Hal-hal tersebut yang mendorong untuk melakukan penelitian ini dengan judul “Implementasi Data Mining Untuk Mengelompokkan Korban Penyalahgunaan Narkoba Menggunakan Metode *K-Means Clustering* Pada BNN Provinsi Sumatera Utara”, diharapkan dengan dibuatnya penelitian ini dapat berguna untuk mengelompokkan korban penyalahgunaan narkoba pada BNN Provinsi Sumatera Utara.

2. METODE PENELITIAN

2.1 Data Mining

Data mining adalah sebuah metode dalam menemukan informasi berharga dari sejumlah data yang dilakukan dengan memanfaatkan ilmu lain seperti statistik, matematika, pengenalan pola [4].

Data Mining (DM) merupakan serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Patut diingat bahwa kata *mining* sendiri berarti usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar. Karena itu *Data Mining* sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik dan database.

Menurut Asmira (2019) data mining memiliki beberapa pengertian yaitu sebagai berikut [5]:

1. Mengekstrak atau *mining* pengetahuan dari kumpulan data yg sangat besar
2. Ekstraksi informasi yang berguna dari data, dimana sebelumnya tidak diharapkan, tidak dikenal dan implisit
3. Eksplorasi dan analisis, secara otomatis atau semi otomatis dari sekumpulan data yg sangat besar untuk memperoleh pola 2 data yg berarti
4. Proses analisis database yang besar secara semi otomatis untuk menemukan pola yang valid, baru, berguna dan dapat dipahami manusia.

Dengan arti lain *Data mining* adalah proses untuk penggalian pola-pola dari data. *Data mining* menjadi alat yang semakin penting untuk mengubah data tersebut menjadi informasi. Hal ini sering digunakan dalam berbagai praktek profil, seperti pemasaran, pengawasan, penipuan deteksi dan penemuan ilmiah. Alasan utama untuk menggunakan data *mining* adalah untuk membantu dalam analisis koleksi pengamatan perilaku. Data tersebut rentan terhadap collinearity karena diketahui keterkaitan. Fakta yang tak terelakkan data *mining* adalah bahwa subset/set data yang dianalisis mungkin tidak mewakili seluruh domain, dan karenanya tidak boleh berisi contoh-contoh hubungan kritis tertentu dan perilaku yang ada di bagian lain dari domain. Untuk mengatasi masalah semacam ini, analisis dapat ditambah menggunakan berbasis percobaan dan pendekatan lain, seperti *Choice Modelling* untuk data yang dihasilkan manusia. Dalam situasi ini, yang melekat dapat berupa korelasi dikontrol untuk, atau dihapus sama sekali, selama konstruksi desain eksperimental. Beberapa teknik yang sering disebut-sebut dalam literatur *Data Mining* dalam penerapannya antara lain: *clustering*, *classification*, *association rule mining*, *neural network*, *genetic algorithm* dan lain-lain.

Data Mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu [8] :

1. Deskripsi
Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.
2. Estimasi
Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan baris data (record) lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.
3. Prediksi
Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.
4. Klasifikasi
Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.
5. *Clustering*

Pengklasteran merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas obyek-obyek yang memiliki kemiripan. Klaster adalah kumpulan record yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan record dalam klaster yang lain. Berbeda dengan klasifikasi, pada pengklasteran tidak ada variabel target. Pengklasteran tidak melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target, akan tetapi, algoritma pengklasteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan record dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan record dalam kelompok lain akan bernilai minimal.

6. Asosiasi

Tugas asosiasi dalam data mining adalah untuk menemukan atribut yang muncul dalam satu waktu. Salah satu implementasi dari asosiasi adalah market basket analysis atau analisis keranjang belanja.

Data Mining bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan *Data Mining* adalah kenyataan bahwa *Data Mining* mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dahulu. Berawal dari beberapa disiplin ilmu, *Data Mining* bertujuan untuk memperbaiki teknik tradisional sehingga bisa menangani [9]:

1. Jumlah data yang sangat besar
2. Dimensi data yang tinggi
3. Data yang heterogen dan berbeda sifat

2.2 K-Means Clustering

K-Means merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data kedalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah meminimalkan variasi didalam suatu kelompok dan memaksimalkan variasi antar kelompok [11].

Pada dasarnya *clustering* merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan data yang lain. *Clustering* merupakan salah satu metode *data mining* yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*taining*) dan tanpa ada guru (*teacher*) serta tidak memerlukan target output. Dalam data mining ada dua jenis metode clustering yang digunakan dalam pengelompokan data, yaitu *hierarchical clustering* dan *non-hierarchical clustering*.

Hierarchical clustering adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan kedua. Demikian seterusnya sehingga cluster akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah *cluster*. Dendrogram biasanya digunakan untuk membantu memperjelas proses heirarki tersebut.

Berbeda dengan metode *hierarchical clustering*, metode *non-hierarchical clustering* justru dimulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan (dua cluster, tiga cluster, atau lain sebagainya). Setelah jumlah cluster diketahui, baru proses cluster dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan *K-Means Clustering*.

Adapun langkah-langkah melakukan clustering dengan metode *K-Means* adalah sebagai berikut [12]:

1. Menentukan banyaknya cluster (k) untuk jumlah cluster dari dataset yang ada.
2. Menentukan k sebagai Centroid, biasanya dilakukan secara acak (random).
3. Hitung jarak data dengan centroid menggunakan rumus jarak menggunakan rumus Euclidean

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

$D(i, j)$ = Jarak data ke i ke pusat cluster j

X_{ki} = Data ke i pada atribut data ke k

X_{kj} = Titik pusat ke j pada atribut ke k

4. Kelompokkan data berdasarkan kedekatan dengan centroid kemudian perbaharui nilai centroid baru dengan lokasi dari pusat cluster menggunakan persamaan.

$$\mu_j(t+1) = \frac{1}{N_{sj}} \sum_{j \in S_j} x_j$$

Dimana $\mu_j(t+1)$ merupakan centroid baru pada iterasi ke (t+1) dan N_{sj} merupakan banyaknya data pada cluster s_j .

5. Lakukan langkah 2 sampai 4 sampai anggota tiap cluster tidak ada yang berubah.

2.3 Unified Modeling Language (UML)

UML (*Unified Modelling Language*) diagram memiliki tujuan utama untuk membantu tim pengembangan proyek berkomunikasi, mengeksplorasi potensi desain, dan memvalidasi desain arsitektur perangkat lunak atau pembuat program. Komponen atau notasi UML diturunkan dari 3 (tiga) notasi yang telah ada sebelumnya yaitu

Grady Booch, OOD (*Object- Oriented Design*), Jim Rumbaugh, OMT (*Object Modelling Technique*), dan Ivar Jacobson OOSE (*Object-Oriented Software Engineering*).

UML (*Unified Modelling Language*) adalah suatu alat untuk memvisualisasikan dan mendokumentasikan hasil analisa dan desain yang berisi sintak dalam memodelkan sistem secara visual. Juga merupakan satu kumpulan konvensi pemodelan yang digunakan untuk menentukan atau menggambarkan sebuah sistem *software* yang terkait dengan objek .

3. ANALISIS DAN HASIL

3.1 Analisis

Dalam menentukan pengelompokkan korban penyalahgunaan narkoba digunakan beberapa jenis data diantaranya yaitu data variabel, data primer dari instansi dan data hasil inisialisasi. Dalam aplikasi data mining menentukan pengelompokkan korban penyalahgunaan narkoba, maka harus ditetapkan variabel-variabel yang digunakan sebagai acuan untuk penilaian dalam proses pengujian. Variabel-variabel tersebut dapat dilihat pada tabel di bawah ini :

Tabel 3.2 Variabel Yang Digunakan

No	Kode	Variabel	Keterangan
1	V1	Usia	Usia korban
2	V2	Zat	Zat yang digunakan
3	V3	Lama Rehab	Lama rehabilitasi

Agar data pada tabel 3.1 di atas dapat diolah dengan menggunakan metode k-means clustering, maka data yang berjenis data bukan nominal seperti Zat dan Lama Rehab harus diinisialisasikan terlebih dahulu dalam bentuk angka. Untuk melakukan inisialisasi atribut dilakukan dengan langkah berikut:

1. Melakukan proses inisialisasi terhadap data Zat guna kepentingan proses clustering. Dengan cara data Zat yang memiliki frekuensi terbesar diberi inisial dengan angka 1 dan Zat yang memiliki frekuensi terbesar kedua diberi inisial dengan angka 2, begitu seterusnya hingga Zat dengan frekuensi paling sedikit. Hasil dari inisialisasi Zat dapat dilihat pada tabel 3.3.

Tabel 3.3 Inisialisasi Data Zat

Zat	Frekuensi	Inisialisasi
Shabu	40	1
Ekstasi	8	2
Inex	6	3

Tabel 3.3 Inisialisasi Data Zat (lanjutan)

Zat	Frekuensi	Inisialisasi
Ganja	3	4
Methamphetamin	3	5

2. Melakukan proses inisialisasi terhadap data Lama Rehab guna kepentingan proses clustering. Dengan cara seperti tabel di atas. Hasil dari inisialisasi data Lama Rehab dapat dilihat pada tabel di bawah ini.

Tabel 3.4 Inisialisasi Data Lama Rehab

Lama Rehab	Frekuensi	Inisialisasi
1 bulan	24	1
< 1 bulan	16	2
2 bulan	15	3
3 bulan	3	4
> 5 bulan	2	5

Sehingga menghasilkan tabel hasil inisialisasi seperti dijelaskan pada tabel di bawah ini.

Tabel 3.5 Tabel Hasil Inisialisasi Data Korban

No	No RM	Alamat	V1	V2	V3
1	2019.12.00.0001	Kec. Medan Johor	39	1	1
2	2019.12.00.0002	Kec. Medan Timur	16	1	1
3	2019.12.00.0003	Kec. Medan Baru	32	1	2
4	2019.12.00.0004	Kec. Medan Baru	31	1	5
5	2019.12.00.0005	Kec. Medan Barat	25	5	1

No	No RM	Alamat	V1	V2	V3
6	2019.12.00.0006	kec. Medan Timur	47	5	1
7	2019.12.00.0007	kec. Medan Timur	21	4	2
8	2019.12.00.0008	Kec. Hamparan Perak	37	1	1
9	2019.12.00.0009	Kec. Rambutan	36	3	1
10	2019.12.00.0010	Kec. Percut Sei Tuan	18	3	3
11	2019.12.00.0011	Kec. Medan Tembung	21	1	3
12	2019.12.00.0012	Kec. Percut Sei Tuan	27	1	3
13	2019.12.00.0013	Kec. Medan Deli	22	1	1
14	2019.12.00.0014	Kec. Medan Deli	22	1	1
15	2019.12.00.0015	kec. Medan Timur	37	1	1
16	2019.12.00.0016	Kec. Medan Area	22	2	4
17	2019.12.00.0017	Kec. Medan Area	23	5	4
18	2019.12.00.0018	Kec. Medan Denai	22	2	1
19	2019.12.00.0019	Kec. Medan Denai	24	2	1
20	2019.12.00.0020	Kec. Pagar Merbau	29	1	2
21	2019.12.00.0021	Kec. Stabat	50	1	3
22	2019.12.00.0022	kec. Medan Timur	50	1	3
23	2019.12.00.0023	Kec. Deli Serdang	46	1	3
24	2019.12.00.0024	Kec. Medan Labuhan	24	1	1
25	2019.12.00.0025	Kec. Batang Kuis	18	1	1
26	2019.12.00.0026	Kec. Medan Perjuangan	42	1	3
27	2019.12.00.0027	Kec. Tanjung Morawa	16	1	3
28	2019.12.00.0028	Kec. Medan Tembung	35	1	1
29	2019.12.00.0029	Kec. Labuhan	25	3	2
30	2019.12.00.0030	Kec. Tanjung Morawa	29	1	3
31	2019.12.00.0031	Kec. Purwolinggo	31	2	2
32	2019.12.00.0032	Kec. Pontianak Timur	34	2	2
33	2019.12.00.0033	Kec. Medan Perjuangan	37	1	1
34	2019.12.00.0034	Kec. Sagulung	39	1	1
35	2019.12.00.0035	Kec. Medan Selayang	45	1	1
36	2019.12.00.0036	Kec. Petisah	35	1	1
37	2019.12.00.0037	Kec. Maimun	23	3	3
38	2019.12.00.0038	Kec. Percut Sei Tuan	24	2	3
39	2019.12.00.0039	Kec. Percut Sei Tuan	33	1	2
40	2019.12.00.0040	Kec. Percut Sei Tuan	20	1	1
41	2019.12.00.0041	Kec. Badar	24	3	3
42	2019.12.00.0042	Kec. Medan Belawan	30	1	1
43	2019.12.00.0043	Kec. Binjai Utara	28	1	1
44	2019.12.00.0044	Kec. Medan Labuhan	45	1	4
45	2019.12.00.0045	Kec. Medan Polonia	24	1	3
46	2019.12.00.0046	Kec. Medan Polonia	24	1	1
47	2019.12.00.0047	Kec. Sunggal	26	1	2
48	2019.12.00.0048	Kec. Medan Marelan	31	1	2
49	2019.12.00.0049	Kec. Medan Marelan	34	4	1
50	2019.12.00.0050	Kec. Medan Labuhan	21	1	2
51	2019.12.00.0051	Kec. Medan Denai	24	2	2
52	2019.12.00.0052	Kec. Simpang Empat	23	1	3
53	2019.12.00.0053	Kec. Medan Kota	38	2	1

No	No RM	Alamat	V1	V2	V3
54	2019.12.00.0054	Kec. Percut Sei Tuan	21	1	2
55	2019.12.00.0055	Kec. Medan Denai	28	1	2
56	2019.12.00.0056	Kec. Medan Deli	21	3	5
57	2019.12.00.0057	Kec. Tanjung Morawa	16	1	3
58	2019.12.00.0058	Kec. Tanjung Morawa	20	1	2
59	2019.12.00.0059	Kec. Medan Denai	35	1	2
60	2019.12.00.0060	Kec. Medan Amplas	40	4	2

Untuk menentukan pusat (centroid) awal pada iterasi ke-1 ini ditentukan dengan acak (random) dari data yang sudah ada. Pada kasus ini pusat centroid awal adalah data ke-6, data ke-30 dan data ke-68:

Tabel 3.6 Titik Pusat (Centroid) Awal Cluster

Centroid	Kode Korban	Alamat	V1	V2	V3
Centroid 1	07	Kec. Medan Timur	21	4	2
Centroid 2	31	Kec. Purwolinggo	31	2	2
Centroid 3	36	Kec. Petisah	35	1	1

Perhitungan jarak dari data ke-1 terhadap titik pusat cluster adalah sebagai berikut dibawah ini:

$$D(1,1) = \sqrt{(39 - 21)^2 + (1 - 4)^2 + (1 - 2)^2}$$

$$= 18,28$$

$$D(1,2) = \sqrt{(39 - 31)^2 + (1 - 2)^2 + (1 - 2)^2}$$

$$= 8,12$$

$$D(1,3) = \sqrt{(39 - 35)^2 + (1 - 1)^2 + (1 - 1)^2}$$

$$= 4$$

Dan seterusnya dilakukan perhitungan jarak untuk data ke-2 sampai data ke-60 dengan rumus seperti diatas. Sehingga akan didapatkan hasil perhitungan jarak setiap data terhadap pusat cluster baru sebagai berikut dijelaskan pada tabel di bawah ini:

Tabel 3.7 Hasil Perhitungan Distance Score

Kode Korban	Alamat	V1	V2	V3	Cluster
1	Kec. Medan Johor	18,28	8,12	4,00	3
2	kec. Medan Timur	5,92	15,07	19,00	1
3	Kec. Medan Baru	11,40	1,41	3,16	2
4	Kec. Medan Baru	10,86	3,16	5,66	2
5	Kec. Medan Barat	4,24	6,78	10,77	1
6	kec. Medan Timur	26,04	16,31	12,65	3
7	kec. Medan Timur	0,00	10,20	14,35	1
8	Kec. Hamparan Perak	16,31	6,16	2,00	3
9	Kec. Rambutan	15,07	5,20	2,24	3
10	Kec. Percut Sei Tuan	3,32	13,08	17,23	1
11	Kec. Medan Tembung	3,16	10,10	14,14	1
Kode Korban	Alamat	V1	V2	V3	Cluster
12	Kec. Percut Sei Tuan	6,78	4,24	8,25	2

13	Kec. Medan Deli	3,32	9,11	13,00	1
14	Kec. Medan Deli	3,32	9,11	13,00	1
15	kec. Medan Timur	16,31	6,16	2,00	3
16	Kec. Medan Area	3,00	9,22	13,38	1
17	Kec. Medan Area	3,00	8,77	13,00	1
18	Kec. Medan Denai	2,45	9,06	13,04	1
19	Kec. Medan Denai	3,74	7,07	11,05	1
20	Kec. Pagar Merbau	8,54	2,24	6,08	2
...
60	Kec. Medan Amplas	19,00	9,22	5,92	3

Setelah dilakukan perhitungan pusat centroid maka selanjutnya menugaskan kembali setiap objek/data dengan pusat cluster baru. Dengan cara melakukan perhitungan jarak pada iterasi ke-2 sebagai berikut :

a. Perhitungan jarak dari data ke-1 terhadap pusat cluster adalah:

$$D(1,1) = \sqrt{(39 - 28,571)^2 + (1 - 1,928)^2 + (1 - 2,214)^2}$$

$$= 10,54$$

$$D(1,2) = \sqrt{(39 - 28,333)^2 + (1 - 1,083)^2 + (1 - 2,25)^2}$$

$$= 10,74$$

$$D(1,3) = \sqrt{(39 - 30,7)^2 + (1 - 1,7)^2 + (1 - 1,7)^2}$$

$$= 8,359$$

Dan seterusnya dilakukan perhitungan jarak untuk data ke-2 sampai data ke-60 dengan rumus seperti diatas. Sehingga akan didapatkan hasil perhitungan jarak setiap data terhadap pusat cluster baru sebagai berikut :

1. Cluster 1 merupakan cluster terkecil dengan anggota cluster 5 data. Mayoritas zat yang digunakan adalah Inex dan Methamphetamin dengan rentang usia korban 21-25 tahun.

Tabel 3.18 Anggota Cluster 1

Kode	No RM	Alamat	Usia	Zat	Lama Rehab
5	2019.12.00.0005	Kec. Medan Barat	25	Methamphetamin	1 bulan
7	2019.12.00.0007	kec. Medan Timur	21	Ganja	< 1 bulan
17	2019.12.00.0017	Kec. Medan Area	23	Methamphetamin	3 bulan
29	2019.12.00.0029	Kec. Labuhan	25	Inex	< 1 bulan
41	2019.12.00.0041	Kec. Badar	24	Inex	2 bulan

2. Cluster 2 merupakan cluster terbesar dengan anggota cluster 29 data. Mayoritas zat yang digunakan adalah Shabu dengan rentang usia korban penyalahgunaan narkoba 16-29 tahun.

Tabel 3.19 Anggota Cluster 2

Kode	No RM	Alamat	Usia	Zat	Lama Rehab
2	2019.12.00.0002	kec. Medan Timur	16	Shabu	1 bulan
10	2019.12.00.0010	Kec. Percut Sei Tuan	18	Inex	2 bulan
11	2019.12.00.0011	Kec. Medan Tembung	21	Shabu	2 bulan
12	2019.12.00.0012	Kec. Percut Sei Tuan	27	Shabu	2 bulan
13	2019.12.00.0013	Kec. Medan Deli	22	Shabu	1 bulan
14	2019.12.00.0014	Kec. Medan Deli	22	Shabu	1 bulan
16	2019.12.00.0016	Kec. Medan Area	22	Ekstasi	3 bulan
18	2019.12.00.0018	Kec. Medan Denai	22	Ekstasi	1 bulan
19	2019.12.00.0019	Kec. Medan Denai	24	Ekstasi	1 bulan
20	2019.12.00.0020	Kec. Pagar Merbau	29	Shabu	< 1 bulan
24	2019.12.00.0024	Kec. Medan Labuhan	24	Shabu	1 bulan
25	2019.12.00.0025	Kec. Batang Kuis	18	Shabu	1 bulan

Kode	No RM	Alamat	Usia	Zat	Lama Rehab
27	2019.12.00.0027	Kec. Tanjung Morawa	16	Shabu	2 bulan
30	2019.12.00.0030	Kec. Tanjung Morawa	29	Shabu	2 bulan
37	2019.12.00.0037	Kec. Maimun	23	Inex	2 bulan
38	2019.12.00.0038	Kec. Percut Sei Tuan	24	Ekstasi	2 bulan
40	2019.12.00.0040	Kec. Percut Sei Tuan	20	Shabu	1 bulan
43	2019.12.00.0043	Kec. Binjai Utara	28	Shabu	1 bulan
50	2019.12.00.0050	Kec. Medan Labuhan	21	Shabu	< 1 bulan
51	2019.12.00.0051	Kec. Medan Denai	24	Ekstasi	< 1 bulan
52	2019.12.00.0052	Kec. Simpang Empat	23	Shabu	2 bulan
54	2019.12.00.0054	Kec. Percut Sei Tuan	21	Shabu	< 1 bulan
55	2019.12.00.0055	Kec. Medan Denai	28	Shabu	< 1 bulan
56	2019.12.00.0056	Kec. Medan Deli	21	Inex	> 5 bulan
57	2019.12.00.0057	Kec. Tanjung Morawa	16	Shabu	2 bulan
58	2019.12.00.0058	Kec. Tanjung Morawa	20	Shabu	< 1 bulan

3. Cluster 3 merupakan cluster dengan anggota cluster 26 data. Mayoritas zat yang digunakan adalah Shabu dengan rentang usia korban penyalahgunaan narkoba 31-50 tahun.

Tabel 3.20 Anggota Cluster 3

Kode	No RM	Alamat	Usia	Zat	Lama Rehab
1	2019.12.00.0001	Kec. Medan Johor	39	Shabu	1 bulan
3	2019.12.00.0003	Kec. Medan Baru	32	Shabu	< 1 bulan
4	2019.12.00.0004	Kec. Medan Baru	31	Shabu	> 5 bulan
6	2019.12.00.0006	kec. Medan Timur	47	Methamphetamin	1 bulan
8	2019.12.00.0008	Kec. Hamparan Perak	37	Shabu	1 bulan
9	2019.12.00.0009	Kec. Rambutan	36	Inex	1 bulan
15	2019.12.00.0015	kec. Medan Timur	37	Shabu	1 bulan
21	2019.12.00.0021	Kec. Stabat	50	Shabu	2 bulan
22	2019.12.00.0022	kec. Medan Timur	50	Shabu	2 bulan
23	2019.12.00.0023	Kec. Deli Serdang	46	Shabu	2 bulan
26	2019.12.00.0026	Kec. Medan Perjuangan	42	Shabu	2 bulan
28	2019.12.00.0028	Kec. Medan Tembung	35	Shabu	1 bulan
31	2019.12.00.0031	Kec. Purwolinggo	31	Ekstasi	< 1 bulan
32	2019.12.00.0032	Kec. Pontianak Timur	34	Ekstasi	< 1 bulan
33	2019.12.00.0033	Kec. Medan Perjuangan	37	Shabu	1 bulan
34	2019.12.00.0034	Kec. Sagulung	39	Shabu	1 bulan
35	2019.12.00.0035	Kec. Medan Selayang	45	Shabu	1 bulan
36	2019.12.00.0036	Kec. Petisah	35	Shabu	1 bulan
39	2019.12.00.0039	Kec. Percut Sei Tuan	33	Shabu	< 1 bulan
42	2019.12.00.0042	Kec. Medan Belawan	30	Shabu	1 bulan
44	2019.12.00.0044	Kec. Medan Labuhan	45	Shabu	3 bulan
48	2019.12.00.0048	Kec. Medan Marelan	31	Shabu	< 1 bulan
49	2019.12.00.0049	Kec. Medan Marelan	34	Ganja	1 bulan
53	2019.12.00.0053	Kec. Medan Kota	38	Ekstasi	1 bulan
59	2019.12.00.0059	Kec. Medan Denai	35	Shabu	< 1 bulan
60	2019.12.00.0060	Kec. Medan Amplas	40	Ganja	< 1 bulan

3.2 Hasil

Implementasi sistem menjelaskan mengenai hasil sistem yang telah dibangun. Terdiri dari beberapa form input dan beberapa laporan. Berikut di bawah ini dijelaskan lebih detail.

1. Form Data Korban

Tampilan ini berisikan tentang data korban yang berfungsi sebagai media dalam memasukan data korban baru dan juga mengedit serta menghapus data korban. Tampilan *form* dirancang agar mudah untuk digunakan oleh *user*. Adapun tampilan *form* sebagai berikut:

Kode Korban	Alamat Korban	Usia	Zat	Lama Rehab
001	Kec. Medan Johor	39	Shabu	1 bulan
002	kec. Medan Timur	16	Shabu	1 bulan
003	Kec. Medan Baru	32	Shabu	< 1 bulan
004	Kec. Medan Baru	31	Shabu	> 5 bulan
005	Kec. Medan Barat	25	Methampetamin	1 bulan
006	kec. Medan Timur	47	Methampetamin	1 bulan
007	kec. Medan Timur	21	Ganja	< 1 bulan
008	Kec. Hamparan Perak	37	Shabu	1 bulan
009	Kec. Rambutan	36	Inav	1 bulan

Gambar Form Data Korban

2. Form Proses Titik Cluster

Tampilan *Form* Proses Titik *Cluster* ini berfungsi untuk menampilkan hasil inialisasi dan memilih 3 data sebagai titik pusat *cluster*. Adapun cara penggunaannya dengan terlebih dahulu memilih (*double click*) pada *listview* pertama maka titik pusat *cluster* akan tampil pada *listview* kedua. Tampilan *form* dapat dilihat pada gambar berikut.

Kode Korban	Alamat Korban	Usia	Zat	Lama Rehab
031	Kec. Purwalinggo	31	2	2
032	Kec. Pontianak Timur	34	2	2
033	Kec. Medan Perjuangan	37	1	1
034	Kec. Sagulung	39	1	1
035	Kec. Medan Selayang	45	1	1
036	Kec. Petaisah	35	1	1
037	Kec. Maimun	23	3	3
038	Kec. Percut Sei Tuan	24	2	3
039	Kec. Percut Sei Tuan	33	1	2

Kode Korban	Alamat Korban	Usia	Zat	Lama Rehab
007	kec. Medan Timur	21	4	2
031	Kec. Purwalinggo	31	2	2
036	Kec. Petaisah	35	1	1

Gambar Tampilan Form Proses Titik Cluster

Adapun fungsi-fungsi dari tombol yang terdapat dalam *form* yaitu :

- Simpan : Menyimpan data Proses Titik *Cluster* baru
- Batal : Membatalkan proses data dan membersihkan *form*

3. Form Proses Clustering

Tampilan *Form Proses Clustering* ini berfungsi untuk melakukan proses perhitungan jarak data ke titik *cluster* menggunakan metode *K-Means Clustering* dan menampilkan hasil perhitungan. Adapun hasil perhitungannya tampil dalam bentuk listview. Klik tombol *Proses* untuk memulai perhitungan dengan metode *K-Means Clustering*. Tampilan *form* sebagai berikut :

Cluster 1

Kode Korban	Alamat Korban	Jarak Cluster 1	Jarak Cluster 2	Jarak Cluster 3	Cluster
002	kec. Medan Timur	5,7773	13,641	24,0704	1
005	Kec. Medan Barat	4,7606	6,1571	15,4234	1
007	kec. Medan Timur	2,1595	9,0688	19,1907	1
010	Kec. Percut Sei Tuan	3,8106	11,7648	22,1265	1
011	Kec. Medan Tembung	1,3439	8,6164	19,1071	1

Cluster 2

Kode Korban	Alamat Korban	Jarak Cluster 1	Jarak Cluster 2	Jarak Cluster 3	Cluster
003	Kec. Medan Baru	10,472	2,431	8,0859	2
004	Kec. Medan Baru	9,8752	3,0946	9,6583	2
012	Kec. Percut Sei Tuan	5,5632	2,6913	13,1333	2
020	Kec. Pagar Merbau	7,4894	0,6401	11,0762	2
030	Kec. Tanjung Morawa	7,5275	0,9538	11,1482	2

Cluster 3

Kode Korban	Alamat Korban	Jarak Cluster 1	Jarak Cluster 2	Jarak Cluster 3	Cluster
001	Kec. Medan Johor	17,4955	9,4996	1,4431	3
006	kec. Medan Timur	25,6422	17,8953	7,7254	3
008	Kec. Hampan Perak	15,5041	7,5217	3,2066	3
009	Kec. Rambutan	14,5192	6,8125	4,3107	3
015	kec. Medan Timur	15,6241	7,5217	3,2066	3

Gambar Tampilan Hasil *Form Proses Clustering*

4. Laporan

Pengujian program yang dilakukan menghasilkan output berupa laporan hasil analisa. *Form* Laporan ini berfungsi untuk melihat hasil perhitungan *Cluster* dengan metode *K-Means Clustering* beserta informasi lain mengenai Titik *Cluster* tersebut. Adapun hasil *Cluster* akan tampil pada kolom *Cluster*. Tampilan *preview* dapat dilihat dibawah ini:

BADAN NARKOTIKA NASIONAL (B N N)
PROVINSI SUMATERA UTARA

Laporan Pengelompokkan Korban Penyalahgunaan Narkoba

No	Kode Korban	Alamat	Jarak Cluster 1	Jarak Cluster 1	Jarak Cluster 3	Cluster Terpilih
1	045	Kec. Medan Polonia	2,7162	5,6341	16,1178	1
2	025	Kec. Batang Kuis	3,8848	11,6509	22,0722	1
3	024	Kec. Medan Labuhan	2,8696	5,7222	16,0805	1
4	019	Kec. Medan Denai	2,7162	5,7945	16,0681	1
5	018	Kec. Medan Denai	1,2897	7,7401	18,0661	1
6	017	Kec. Medan Area	3,8293	7,8577	17,5181	1
7	016	Kec. Medan Area	1,8378	7,8364	18,1984	1
8	037	Kec. Maimun	1,9509	6,8976	17,1488	1
9	038	Kec. Percut Sei Tuan	2,5535	5,7076	16,1054	1
10	014	Kec. Medan Deli	1,5876	7,6861	18,0771	1
11	040	Kec. Percut Sei Tuan	2,1923	9,6649	20,0744	1
12	041	Kec. Badar	2,7683	5,9506	16,1550	1
13	013	Kec. Medan Deli	1,5876	7,6861	18,0771	1
14	027	Kec. Tanjung Morawa	5,7027	13,6043	24,0953	1
15	011	Kec. Medan Tembung	1,3439	8,6164	19,1071	1
16	052	Kec. Simpang Empat	1,8763	6,6264	17,1138	1
17	002	kec. Medan Timur	5,7773	13,6410	24,0704	1
18	058	Kec. Tanjung Morawa	1,8378	9,5870	20,0645	1
19	057	Kec. Tanjung Morawa	5,7027	13,6043	24,0953	1
20	056	Kec. Medan Deli	3,0389	9,2146	19,3774	1

Gambar Laporan Hasil *Clustering*

4. KESIMPULAN

Berdasarkan pembahasan dan evaluasi dari bab terdahulu, maka dapat ditarik beberapa kesimpulan sebagai berikut :

1. Dalam pengelompokkan korban penyalahgunaan narkoba di BNN Provinsi Sumatera Utara dilakukan dengan menganalisa data mentah dari BNN Provinsi Sumatera Utara kemudian data tersebut diinisialisasikan kedalam bentuk angka lalu dilakukan proses perhitungan *distance score* sehingga terbentuk hasil *cluster*.
2. Dalam merancang aplikasi data mining yang mengadopsi algoritma *K-Means Clustering* maka diperlukan sebuah bahasa pemrograman *Visual Basic .Net 2008* serta sebuah aplikasi database *Microsoft Access*. Dan dalam pemodelan sistem menggunakan pemodelan *Unified Modeling Language (UML)*.
3. Penerapan metode *K-Means Clustering* dilakukan dengan mengintegrasikan algoritma *K-Means Clustering* kedalam baris kode program *Visual Basic .Net 2008*. Setelah setiap tahapan dalam metode *K-Means Clustering* diletakkan pada baris kode program maka diuji sampai hasil perhitungan *clustering* sesuai dengan hasil perhitungan manual.

REFERENSI

- [1] C. A. Sugianto, A. H. Rahayu, and A. Gusman, "Algoritma K-Means Untuk Pengelompokkan Penyakit Pasien Pada Puskesmas Cigugur Tengah," no. August, 2020, doi: 10.47292/joint.v2i2.30
- [2] Wirdah Choiriah, "Penggunaan Algoritma Apriori Data Mining Untuk Mengetahui Tingkat Kesetiaan Konsumen (*Brand Loyalty*) Terhadap Merek Kendaraan Bermotor (Studi Kasus Dealer Honda Rumbai)," no. Februari, 2016.
- [3] Baginda Harahap, "Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)," 2019.
- [4] E. Iswandy, "Analisa Dan Perancangan Sistem Informasi Penagihan Purchasing Order Customer Studi Kasus Pada Cv. Vertical Cipta Relasi Padang Dengan Metode Centralized Data Processing," *J. TEKNOIF*, vol. 4, no. Oktober, p. 14, 2016.
- [5] R. Yanto and R. Khoiriah, "Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat," pp. 102–113.
- [6] A. Setiawan, et al, "Rancang Bangun Edugame The World of Word Berbasis unity 3D Dengan Implementasi Speech Recognition" *Jurnal Teknik Informatika*, 2017.
- [7] A. Rifa'i, G. Guntoero Setiaji, and V. Vydia, "PENGUNAAN METODE K-MEANS PADA ANALISA DAN KLASIFIKASI CAPRES 2019 DI TWITTER," *Pengemb. Rekayasa dan Teknol.*, vol. 15, no. 1, pp. 43–47, 2019, [Online]. Available: <http://journals.usm.ac.id/index.php/jprt/index>.
- [8] N. Dwitri, J. A. Tampubolon, S. Prayoga, and P. P. P. A. N. W. F. I. R. H. Zer, "PENERAPAN ALGORITMA K-MEANS DALAM MENENTUKAN TINGKAT PENYEBARAN PANDEMI COVID-19 DI INDONESIA," vol. 4, no. 1, pp. 128–132, 2020.
- [9] F. Indriyani and E. Irfiani, "Clustering Data Penjualan Pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means," 2019.
- [10] Indra Griha and George Pri Hartawan, "Perancangan Aplikasi Koperasi Simpan Pinjam berbasis Web", *Jurnal Ilmiah Ilmu Ekonomi*, 2017.
- [11] E. Affandi and T. Syahputra STMIK Triguna Dharma, "J-SISKO TECH Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD Pemodelan Uml Manajemen Sistem Inventory," vol. 1, no. 2, pp. 14–25, 2018.

BIOGRAFI PENULIS

	<p>Yanuar Ishaq Hasibuan, pria kelahiran Siabu, 23 Januari 1996 ini merupakan seorang mahasiswa yang sedang menempuh pendidikan tingkat akhir di STMIK Triguna Dharma Medan jurusan Sistem informasi stanbuk 2015. Bidang riset yang ditekuni adalah BNN Sumatera Utara. Beliau merupakan anak keenam dari bapak Martua hasibuan dan ibu Sumirlak harahap. Rekam pendidikannya yaitu , SDN 07 siabu, SMP Negeri 1 siabu dan Sma Negeri 1 siabu. Saat ini sedang berjuang untuk mengerjakan skripsi guna untuk syarat kelulusan strata 1. Dengan mengangkat sebuah judul “Implementasi Data Mining Untuk Mengelompokkan Korban Penyalahgunaan Narkoba Menggunakan Metode <i>K-Means Clustering</i> Pada BNN Provinsi Sumatera Utara ”.</p>
	<p>Puji Sari Ramdhan, S.Kom M.Kom, jabatan Kaprodi sistem informasi Triguna Dharma medan Sumatera Utara....</p>
	<p>Milfa Yetri, S.Kom., M.Kom perempuan ini merupakan kelahiran Medan, 09 Maret 1988,, mata kuliah yang diampaiu :Teknik Presentasi, Arsitektur Komputer ,Animasi, Pengantar Teknologi Informasi, Perakitan dan Perawatan KomputerData,Mining. Beliau saat ini merupakan dosen tetap di STMIK Triguna Dharma Medan. Alamat kantor Jl. Pintu Air1/Jend. A.H Nasution No.73 Medan Johor. Tamat Strata 1 Sistem Informasi STMIK Triguna Dharma dan Strata 2 Magister Ilmu Teknologi informasi YPTK Padang.</p>